ACOUSTIC ANALYSIS OF EMOTIONAL SPEECH PROCESSED BY HEARING AIDS

Huiwen Goy*[1], M. Kathleen Pichora-Fuller†[2], Gurjit Singh‡[3,4] and Frank A. Russo#[1]
[1]Ryerson University, Toronto, Ontario, Canada.
[2]University of Toronto, Mississauga, Ontario, Canada.
[3]Toronto Rehabilitation Institute, Toronto, Ontario, Canada.
[4]Phonak Canada, Mississauga, Ontario, Canada.

1 Introduction

The speech signal carries important information about the emotional state of the talker. Listeners with hearing loss experience difficulties identifying vocal emotion, possibly due to threshold hearing loss [1] and/or changes in supra-threshold auditory or cognition processing [2]. Some evidence shows that current generation hearing aids do not work well for improving vocal emotion perception [3]. It is likely that the processing of speech sounds to increase audibility also changes the acoustic cues used for emotion perception, including amplitude and spectral cues [4, 5]. Increasing the gain for low-level sounds more than the gain for high-level sounds makes the amplitude envelope of speech less variable. Increasing the amount of high-frequency energy to meet hearing aid fitting targets changes the spectral characteristics of speech. One study found that listeners with hearing aids are in fact more sensitive to intensity cues than normal-hearing listeners on an arousal rating task [6]. Comparing the effects of different types of simulated hearing aid processing on vocal emotional cues may inform us on how current hearing aids affect emotion perception by listeners with hearing loss.

2 Method

2.1 Original speech materials

There were 140 sound files selected from the recordings of the younger female talker in the Toronto Emotional Speech Set [7]. These sound files consisted of 20 different sentences, each spoken in 7 emotion conditions: Angry, disgust, fear, happy, neutral, pleasant surprise and sad. Sentences consisted of the carrier phrase *Say the word* followed by a monosyllabic keyword.

2.2 Hearing aid processing conditions

In the Unaided condition, the recordings were processed to simulate the signal received by a listener with sloping bilateral hearing loss (a pure-tone average of 46.25 dB HL at 0.5, 1, 2 and 4 kHz). In three additional conditions corresponding to three types of hearing aid processing, the recordings were processed using a Phonak hearing aid simulator according to NAL-NL2 targets [8] for the same hearing loss simulated in the Unaided condition.

* huiwen.goy@utoronto.ca
† k.pichora.fuller@utoronto.ca
‡ gurjit.singh@phonak.com
# russo@psych.ryerson.ca

In the Linear condition, the same amount of gain was applied regardless of the sound input level. In Slow Compression and Fast Compression, there was more gain applied for low-intensity than high-intensity sounds, with a compression ratio ranging from 1.1 at lower frequencies to 2.9 at higher frequencies. The speed of compression was about twice as fast in Fast Compression as in Slow Compression. The processed speech was played at 70 dB SPL_A from a loudspeaker in a sound-attenuating booth, and speech was recorded using microphones in the ear canals of a mannequin. All other hearing aid features (directional processing, SoundRecover, etc.) were disabled.

2.3 Acoustical analysis

Since duration and F_0 are not expected to be affected by these processing conditions, only the following measures were taken using the Praat speech analysis program [9]: mean intensity, intensity standard deviation (Intensity SD), and spectral centre-of-gravity (Spectral CoG).

2.4 Statistical analysis

For each acoustic measure, the effects of processing conditions were compared using pairwise *t*-tests with Holm correction. To examine whether emotion conditions were affected differently by processing conditions, an analysis of variance was conducted for Intensity SD and Spectral CoG, with Processing Condition and Emotion as within-subject factors. Significant interactions were analyzed by comparing Emotion conditions within each Processing Condition.

3 Results

3.1 Effects of hearing aid processing on speech acoustic measures

Mean intensity differed across all processing conditions (p's < 0.001), with the lowest overall intensity in the Unaided condition, followed by Fast Compression, Slow Compression and Linear (+3, +8, and +9 dB relative to Unaided, respectively). Intensity SD also differed across processing conditions (p's < 0.001), with the greatest intensity variation in Linear and the least variation in Fast Compression (Figure 1).

Spectral CoG differed across all processing conditions (p's < 0.01), with Fast Compression having the highest Spectral CoG and Unaided having the lowest Spectral CoG (Figure 2).

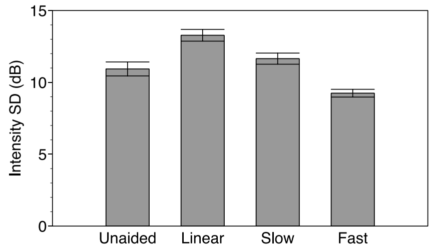

Figure 1: Intensity SD across processing conditions.

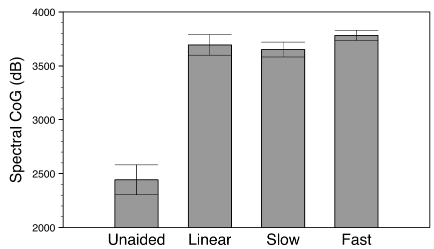

Figure 2: Spectral CoG across processing conditions.

3.2 Interaction of hearing aid processing with emotion condition

There was a main effect of Processing Condition on Intensity SD, $F_{(3, 57)} = 396.3$, $p < 0.001$, a main effect of Emotion, $F_{(6, 114)} = 47.09$, $p < 0.001$, and an interaction of Processing Condition with Emotion, $F_{(18, 342)} = 36.6$, $p < 0.001$. Emotion conditions with the largest intensity variation in the Unaided condition (Angry, Happy) were disproportionately affected by amplitude compression (Figure 3). Some pairs of emotions were no longer distinguishable, e.g., Angry and Disgust were significantly different in the Unaided condition ($p < 0.001$), but not in the Fast Compression condition ($p = 0.54$).

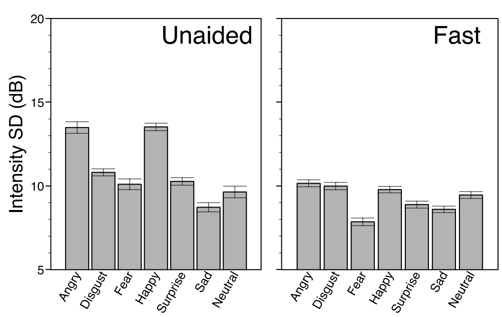

Figure 3: Intensity SD across emotion conditions, in the unaided and fast compression conditions.

There was a main effect of Processing Condition on Spectral CoG, $F_{(3, 57)} = 454$, $p < 0.001$, a main effect of Emotion, $F_{(6, 114)} = 26.23$, $p < 0.001$, and an interaction of Processing Condition with Emotion, $F_{(18, 342)} = 68.14$, $p < 0.001$. The emotion condition with the lowest CoG in the Unaided condition (Sad) became the condition with the highest Spectral CoG in Fast Compression (Figure 4). The emotion condition with a higher CoG than any other emotion condition in the Unaided condition (Angry) became very similar to other emotion conditions in Fast Compression.

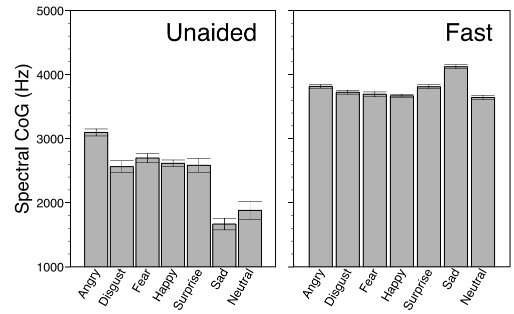

Figure 4: Spectral CoG across emotion conditions, in the unaided and fast compression conditions.

4 Discussion

In this study, simulated hearing aid processing affected two acoustic cues used in emotion perception, namely, intensity variation and spectral cues. The effects of hearing aid processing varied according to the type of vocal emotion. In some cases, hearing aid processing led to emotions being less distinguishable on these two acoustic cues. Future directions may include testing listeners with normal hearing and hearing loss on these processed recordings to determine how changes in specific acoustic cues affect emotion perception.

Acknowledgments

The authors thank Hannes Wuthrich at Phonak for his help with processing sound files.

References

[1] Mitchell RLC, Kingston RA. Age-related decline in emotional prosody discrimination: Acoustic correlates. Exp Psychol. 2014:61(3):215-223.

[2] Mitchell RLC. Age-related decline in the ability to decode emotional prosody: Primary or secondary phenomenon? Cogn Emot. 2007:21(7):1435-1454.

[3] Goy H, Pichora-Fuller MK, Singh G, Russo FA. Perception of emotional speech by listeners with hearing aids. Can Acoust. 2016:44(3):182-183.

[4] Banse R, Scherer KR. Acoustic profiles in vocal emotion expression. J Pers Soc Psychol, 1996:70(3):614-636.

[5] Sobin C, Alpert M. Emotion in speech: The acoustic attributes of fear, anger, sadness, and joy. J Psycholinguist Res. 1999:28(4):347-365.

[6] Schmidt J, Herzog D, Scharenborg O, Janse E. Do hearing aids improve affect perception? Adv Exp Med Biol. 2016:894:47-55.

[7] Dupuis K, Pichora-Fuller MK. Effects of emotional content and emotional voice on speech intelligibility in younger and older adults. Can Acoust. 2008:35(3):114-115.

[8] Keidser G, Dillon H, Flax M, Ching T, Brewer S. (2011). The NAL-NL2 prescription procedure. Audiol Res. 2011:1(e24):88-90.

[9] Boersma P, Weenink D. (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.30.